ELSEVIER

Atmospheric Environment 38 (2004) 4501–4507

ATMOSPHERIC ENVIRONMENT

www.elsevier.com/locate/atmosenv

Identifying the impact of large urban airports on local air quality by nonparametric regression

K.N. Yu[a], Y.P. Cheung[a], T. Cheung[a], Ronald C. Henry[b,*]

[a] *Department of Physics and Materials Science, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong*
[b] *Environmental Engineering Program, Civil Engineering Department, University of Southern California, 3620 South Vermont Avenue, Los Angeles, CA 90089-2531, USA*

Received 4 January 2004; accepted 14 May 2004

Abstract

Emissions of air pollutants from aircraft in large urban areas are a health concern to nearby residents. This study examined hourly concentrations of CO, NO_x, SO_2, and respirable suspended particles (RSP) taken in the vicinity of Hong Kong International Airport (HKIA) and Los Angeles International Airport (LAX). The LAX data cover the period August 1997 through March 1998 and the HKIA all of 2000 and 2001. The average concentration as a function of wind speed and direction was estimated by nonparametric regression. The error variance of the nonparametric regression results was also estimated. The results show that SO_2 can be used to identify wind speeds and directions associated with emissions from aircraft. Using this assumption and the nonparametric regression plots for the other pollutants one can identify the impact of aircraft on local air quality. At LAX, CO and NO_x are dominated by emissions from ground vehicles going in and out of the airport. However, near HKIA, aircraft are an important contributor to CO and RSP. At both sites, nonparametric regression identified other, smaller sources as well.

Keywords: Air pollution; Airports; Hong Kong; Los Angeles; Receptor modeling; Statistics; Data analysis; Nonparametric regression

1. Introduction

The effect of airports on local air quality is of growing concern. In many great cities of the world, emissions from automobiles, power plants, refineries, and other major sources are being steadily reduced. As a result, emissions from aircraft operations, which have stayed the same or increased, are a growing contributor to air pollution, especially in the immediate vicinity of airports. Assessment of the potential impact of airports on the health of nearby residents requires estimates of the contribution of airport operations to pollutant concentrations. This can be done by either traditional source-oriented air quality modeling or by receptor modeling methods, both of which face significant challenges.

Applying standard air quality models to aircraft operations is difficult for several reasons. The emission rates of pollutants from aircraft turbine and piston engines are not well known, and the emissions vary greatly during takeoff, landing, and taxiing, being a maximum during taxi and takeoff (Popp et al., 1999). During takeoff emissions are not a point or line source, but follow a curved path of varying height that is difficult to model realistically. Furthermore, many airports are on the coast with complex winds from lake or sea breezes. Receptor modeling is an alternative to source-oriented models. It uses chemical composition of the emissions to distinguish between sources. In the case of airports, this task is usually made difficult by the presence of a high volume of diesel vehicle emissions

---

*Corresponding author. Tel.: +1-213-740-0596; fax: +1-213-744-1426.
E-mail address: rhenry@usc.edu (R.C. Henry).

1352-2310/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2004.05.034

delivering goods and people to and from the airport. Jet fuel is very similar to diesel fuel (Spicer et al., 1992), making it difficult to distinguish aircraft from diesel emissions using ordinary receptor modeling methods.

The variation of concentrations of pollutants with wind direction and speed is potentially a way of separating out the mix of sources around a large airport. Henry et al. (2002) have shown that nonparametric regression of pollutant concentrations on wind direction is an accurate way to determine the direction of nearby sources. The present paper extends this method to nonparametric regression of hourly pollutant concentrations on two variables, wind direction and wind speed. Nonparametric regression is used to estimate the average concentration of a pollutant such as sulfur dioxide (SO_2) as a function of wind direction and speed. As will be shown below, wind speed is useful in distinguishing ground level emissions from elevated emissions like aircraft. In this way, the contribution of airport operations can be distinguished from vehicular traffic and other sources.

2. Data

Two sets of data are used in this study one from a site very near Los Angeles International Airport (LAX) and one from a site about 3 km from Hong Kong International Airport (HKIA). The LAX data were taken as part of a special study conducted for Los Angeles World Airports; the HKIA data are from a routine air quality monitoring site run by the Hong Kong Environmental Protection Department (HKEPD).

2.1. LAX air quality data

LAX is the third busiest airport in the world as judged by number of passengers passing though it; in other terms, it has about 60,000 aircraft operations (basically takeoffs and landings) a month. The LAX data used in this work was taken as part of a special field study conducted between August 13, 1997 and March 31, 1998. Hourly concentrations of carbon monoxide (CO), (SO_2), nitrogen oxides (NO), NO_2, and meteorological data were made about 200 m from the east end of the south runway (official designation runway 25 right), which services most international and long distance flights, see Fig. 1. LAX is located on the seacoast and aircraft take off to the west over the Pacific Ocean. The monitoring station was directly downwind of the runway (azimuth 263°) and in-line with the path of aircraft taking off and landing. The inlet of the monitors was about 3 m above the ground. The area around LAX is a mixture of residential, commercial and industrial facilities with very heavy traffic from gasoline vehicles and light and heavy-duty diesel vehicles. On the coast, just south of the airport are a small power plant and a major oil refinery (Fig. 1).

Fig. 1. Location of the LAX monitoring site, major roads, and other SO_2 sources.

2.2. HKIA air quality data

The HKEPD currently maintains a network of 14 air-quality monitoring stations for measuring major air pollutants. It consists of 11 stations for monitoring general air quality and 3 stations for roadside air quality across the territory. Most of these stations are located on rooftops from 3 to 25 m above ground. Continuous automatic analyzers are used to measure NO, nitrogen dioxide (NO_2), ozone (O_3), CO, SO_2 and respirable suspended particulates (RSP). Meteorological parameters including temperature, wind speed, wind direction and solar radiation are also recorded continuously at each station. Tung Chung is the monitoring station at the north of Lantau Island and is the closest station to HKIA; it is about 3 km southeast, see Fig. 2. The station monitors general air quality and is located at 21 m above the ground. Tung Chung is a relatively less densely populated area of residential development and light industry.

HKIA has two parallel runways (commonly called the north runway and the south runway), which run northeast to southwest. The HKIA opened in July 1998 with only the South Runway in operation. The numbers of landing and takeoffs were 90,953 and 90,974, respectively, in 2000, and were 98,423 and 98,410, respectively, in 2001.

3. Nonparametric regression methodology

Nonparametric regression is a method of estimating the mean value of a dependent variable given the value of one or more predictor variables (Härdle, 1990; Wand and Jones, 1995). Ordinary regression does the same thing, except that a functional relationship between the

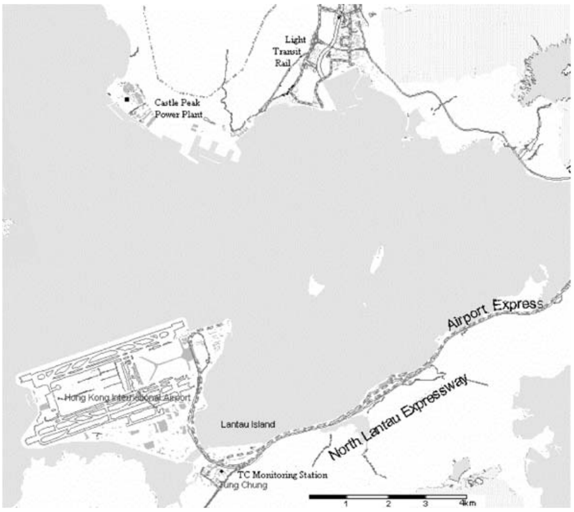

Fig. 2. Location of Tung Chung air quality monitoring station, major roads, and SO_2 sources.

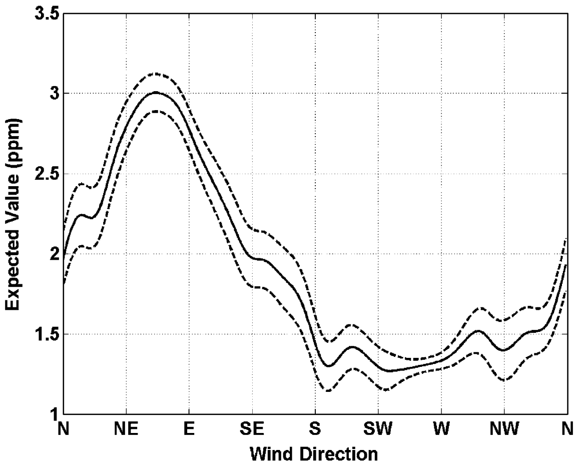

Fig. 3. Nonparametric regression on wind direction of CO for LAX data with wind speed greater that 1 km/h. The dashed lines are 95% confidence intervals. The FWHM of the Gaussian kernel is 26.

dependent and predictor variable is assumed to be known. The most common functional forms are linear, polynomial, exponential, and logarithmic. The data are then used to estimate the unknown parameters of the function; for example, the intercept and slope of a linear function. Thus, ordinary regression can be called parametric regression. Nonparametric regression is a method of estimating the relationship between the dependent and predictor variables without making any assumptions about the functional form of the relationship or the statistical distribution of the data.

In a previous paper (Henry et al., 2002), nonparametric regression was used to estimate the direction to a nearby source of air pollution. An example of nonparametric regression on wind direction alone using CO data from LAX is given in Fig. 3. The highest average concentration of CO in the figure occurs in the direction of the intersection of the two busiest routes to LAX, the San Diego Freeway and Century Boulevard. Although plots with wind direction alone are very useful in determining the direction to nearby sources, there is the problem of travel time and wind speed. If a source is 5 km from the monitoring site, wind speeds greater than or equal to 5 km/h are needed to bring the emissions to the monitor. Also, if a nearby source is elevated above the ground, its plume may pass over the monitor unless the wind speed is great enough to bend it over sufficiently to touch the ground. Clearly, adding wind speed to wind direction in the nonparametric regression of a pollutant can give important new information. This is indeed the case, as the results of this work will show. First, the technical details of nonparametric regression on two explanatory variables are given below.

In this work the dependent variable is the hourly averaged concentration of a pollutant C and the predictor variables are the average wind direction θ and wind speed u during the hour. The average concentration of a pollutant for a particular wind speed and direction pair (θ, u) is calculated as a weighted average of the concentration data in a window around (θ, u) represented by smoothing parameters σ and h using a weighting function $K(\theta, u, \sigma, h) = K_1(\theta, \sigma)K_2(u, h)$, also known as the kernel function. The estimated average concentration is given by

$$\bar{C}(\theta, u) = \frac{\sum_{i=1}^{N} K_1\left(\frac{(\theta - W_i)}{\sigma}\right) K_2\left(\frac{(u - U_i)}{h}\right) C_i}{\sum_{i=1}^{N} K_1\left(\frac{(\theta - W_i)}{\sigma}\right) K_2\left(\frac{(u - U_i)}{h}\right)}, \qquad (1)$$

where C_i U_i, and W_i are the observed concentration of a particular pollutant, resultant wind speed and direction, respectively, for the ith observation in a time period starting at time t_i; N is the total number of observations. Two well-known choices for the kernel functions, and the ones used in this work, are the Gaussian kernel given by

$$K_1(x) = (2\pi)^{-1/2}\exp(-0.5x^2) \quad -\infty < x < \infty \qquad (2)$$

and the Epanechnikov kernel

$$K_2(x) = 0.75(1 - x^2) \quad -1 \leqslant x \leqslant 1. \qquad (3)$$

The Gaussian kernel is used for wind direction since it is defined over an unbounded range, and the Epannechnikov kernel is used for wind speed because it is the simplest bounded kernel.

The smoothing parameters σ and h in Eq. (1) are the only adjustable parameters in nonparametric regression.

To put smoothing parameters for the two different kernels on the same basis, one can use the full-width at half-maximum (FWHM) of the function. The FWHM is easier to understand intuitively than the smoothing parameter, so in this work the smoothing parameter is defined in terms of the FWHM. For a function F symmetric about zero, and with maximum value at zero, FWHM = $2y_1$, where $F(y_1) = 0.5F(0), y_1 > 0$. For the Gaussian kernel, the smoothing parameter σ is related to the FWHM by

$$\sigma = \frac{\text{FWHM}}{2\sqrt{-2\ln(0.5)}}. \tag{4}$$

For the Epanechnikov kernel, the smoothing parameter h is given by

$$h = \frac{\text{FWHM}}{\sqrt{2}}. \tag{5}$$

Methods to estimate values of the smoothing parameters that are optimal under certain assumptions are available (Wand and Jones, 1995). In this work, smoothing parameters were chosen subjectively to give results that were not oversmoothed or too "lumpy". Experience shows that the results are not very sensitive to the choice of the smoothing parameters. For ease of comparison, all the results given below use a FWHM of 25° for the Gaussian kernel and a FWHM of 4 km/h for the Epannechnikov kernel.

Estimates of the uncertainty in the nonparametric regression results can be very useful. For example, air quality data are often subject to outliers, i.e., data that are unusually large. These may be real, unusual values or the result of contamination, analytical errors, data transcription error, or other accidents. Even one large outlier can produce a false peak in nonparametric regression results. However, this peak will have a large uncertainty and can be discounted in the interpretation, if an estimate of the uncertainty at each point is known.

The variance of the average concentration in Eq. (1) is estimated by

$$\operatorname{Var}(\bar{C}(\theta,u)) = \|K_1\|_2^2\|K_2\|_2^2 \frac{\sum_{i=1}^{N} K_1\left(\frac{(\theta - W_i)}{\sigma}\right) K_2\left(\frac{(u - U_i)}{h}\right)(C_i - \bar{C}(\theta,u))^2}{\left[\sum_{i=1}^{N} K_1\left(\frac{(\theta - W_i)}{\sigma}\right) K_2\left(\frac{(u - U_i)}{h}\right)\right]^2}, \tag{6}$$

where

$$\|K_1\|_2^2 = \int_{-\infty}^{\infty} K_1^2(x)\,\mathrm{d}x = 1/2\sqrt{\pi}, \text{ for the Gaussian}$$

kernel and,

$$\|K_2\|_2^2 = \int_{-\infty}^{\infty} K_2^2(x)\,\mathrm{d}x = 0.6, \text{ for the Epanechnikov}$$

kernel.

The above result is derived from more general expressions given in Wand and Jones (1995). Being derived from an asymptotic distribution, this variance estimate is only approximate; but experience has shown it to be quite reasonable, even if there is substantial serial correlation in the data.

4. Results and discussion

4.1. Los Angles international airport

Fig. 4 gives the nonparametric regression of CO concentrations on wind speed and direction for the LAX data. There is a broad area of high concentration associated with winds from approximately 0° to 140° that is a maximum at low wind speeds and decreases as wind speed increases. Closer examination shows that the direction of the maximum in CO concentration points towards the intersection of the San Diego Freeway and Century Blvd., the main routes to the airport. The peak is consistent with what would be expected for CO from ground vehicles. It is very broad covering the region of high traffic volume near the airport. Also, since vehicle emissions are very close to the monitor and emitted at ground level, the concentration decreases with increasing wind speed as expected. However, what is not seen is perhaps more interesting than that what is. Although LAX is one of the busiest airports in the world and aircraft taxiing and taking off emit large amounts of CO, there are no high values of CO around 270°, the direction of the runways. Thus, compared to ground

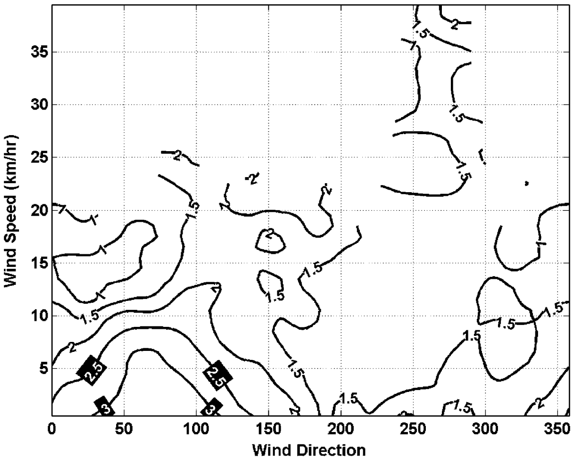

Fig. 4. Nonparametric regression of LAX CO on wind speed and direction. The average CO concentration contours are given in ppm. Contours are not shown at points where the average CO value is less than twice the estimated standard deviation. FWHM for wind direction is 25 and 4 for wind speed. Contours with high values are shown as white numbers on a black background.

vehicle traffic, aircraft do not seem to be a significant source of CO near the airport, at least in terms of producing high concentrations. The nonparametric regression plots for NO, NO_2, and NO_x are almost the same as CO, indicating that ground vehicle traffic dominates these pollutants as well.

The effect of aircraft emissions is clearly seen in Fig. 5, the nonparametric regression of hourly SO_2 concentrations on wind speed and direction. In this figure, the highest concentrations are seen at high wind speeds and around the direction of 270°. While the south runway is oriented at an azimuth of about 263°, the emissions of aircraft from the north runway also affect the monitor. As a result, the maximum concentration of SO_2 is displaced somewhat to the north of 263°. Emissions from aircraft taking off require high winds to bring the emissions down to the ground level. There are relatively high values of SO_2 at lower wind speed between 200° and 300°. These are likely associated with aircraft on the ground and other ground support vehicles. Some of the SO_2 around 220–230° may be associated with the small power plant and the large refinery in this direction. There is some SO_2 associated with the same region of high CO from 0° to 125° and this can be attributed to ground vehicle traffic. Finally, there is evidence of an elevated SO_2 source around 150°. A check of the emission inventory for the area shows a refinery about 10 km distant in the direction 150° with a flare that is a source of SO_2.

The use of SO_2 as a tracer for aircraft emissions is entirely made possible by the ability of nonparametric regression plots such as Fig. 5 to separate aircraft SO_2 from other sources of SO_2. Sulfur is typically 0.05% in Jet A fuel and ordinary diesel fuel (Colls, 2002). The estimated emissions of SO_2 from LAX in 1996 were 210 tons per year (unpublished LAX Master Plan EIS/EIR, revised June, 2000). This is just 1% of the approximately 21,000 tons SO_2 per year released in the whole South Coast Air Basin. This same report estimated the SO_2 from airport support vehicles to be just 2 tons per year, while emissions from cars and trucks coming into the airport was about 27 tons per year. Thus, most of the emissions of SO_2 from LAX are probably associated directly with aircraft, and though these emissions are small compared to the emissions of the whole Los Angeles area, SO_2 can be used as a tracer for the impact of a major airport. This result is confirmed by analysis of SO_2 data from the vicinity of Hong Kong International Airport given later in this paper.

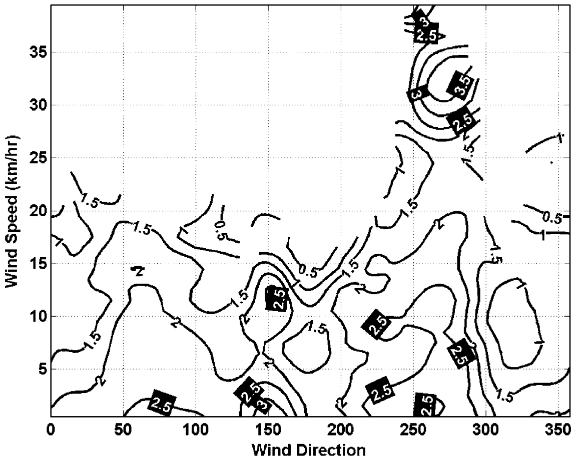

Fig. 5. Nonparametric regression of LAX SO_2 on wind speed and direction. The average SO_2 concentration contours are given in ppb. Contours are not shown at points where the average SO_2 value is less than twice the estimated standard deviation. FWHM for wind direction is 25 and 4 for wind speed. Contours with high values are shown as white numbers on a black background.

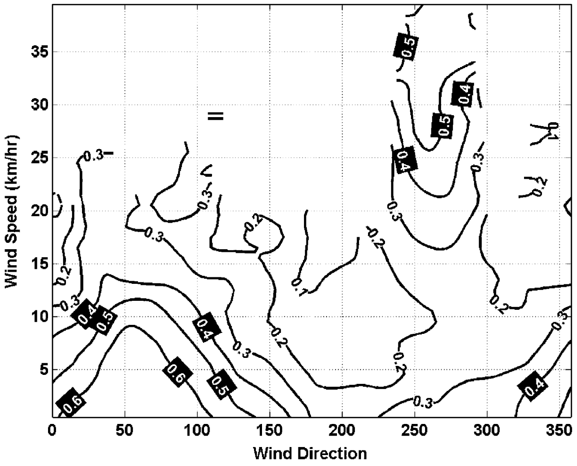

Fig. 6. Nonparametric regression of the LAX NO/NO_x ratio on wind speed and direction. Contours are not shown at points where the average NO/NO_x value is less than twice the estimated standard deviation. FWHM for wind direction is 25 and 4 for wind speed. Contours with high values are shown as white numbers on a black background.

Fig. 6 is the nonparametric regression plot of the ratio of NO to NO_x. This ratio is high for areas affected by fresh combustion emissions, which are mostly NO. The figure clearly shows fresh ground vehicle emissions from 0° to 140°. Fresh emissions are also seen in the region of high SO_2 with high winds and direction 270. This reinforces the association of this SO_2 with aircraft emissions. This association is also seen in the data from the vicinity of HKIA.

4.2. Tung Chung

Fig. 7 is the nonparametric regression plot of hourly SO_2 at Tung Chung versus wind direction and wind

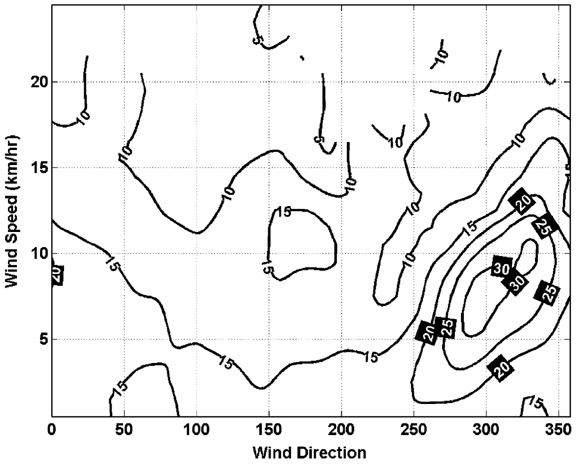

Fig. 7. Nonparametric regression of Tung Chung SO_2 on wind speed and direction. The average SO_2 concentration contours are given in μg/m^3. Contours are not shown at points where the average SO_2 value is less than twice the estimated standard deviation. FWHM for wind direction is 25 and 4 for wind speed. Contours with high values are shown as white numbers on a black background.

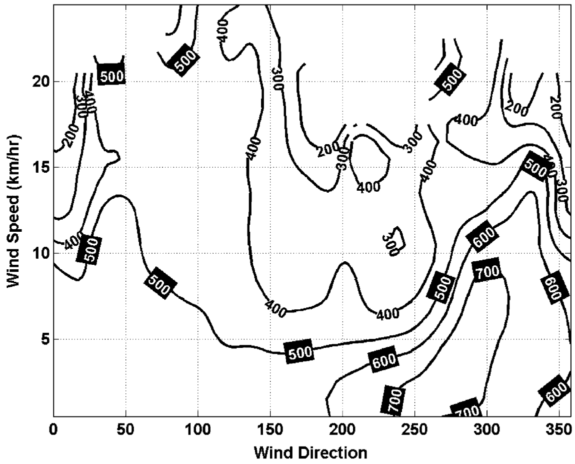

Fig. 8. Nonparametric regression of Tung Chung CO on wind speed and direction. The average CO concentration contours are given in μg/m^3. Contours are not shown at points where the average CO value is less than twice the estimated standard deviation. FWHM for wind direction is 25 and 4 for wind speed. Contours with high values are shown as white numbers on a black background.

speed. There is only one area of high SO_2 from about 250° to 350° and wind direction and between 3 and 13 km/h wind speed. This broad area of high SO_2 is consistent with emissions from aircraft from HKIA. Unlike the LAX monitoring station, the monitoring station at Tung Chung is not located directly inline with the airport runways (Fig. 2). Thus, as seen from Tung Chung, the SO_2 emissions are spread over a wide range of directions, which includes not just the runways but also the takeoff and landing approaches. The dependence on wind speed is also what is expected of emissions from aircraft. Tung Chung is about 3–4 km from the airport so winds of at least 3–4 km/h are needed to bring the emissions to the monitor. Indeed, from the figure high concentrations do not start until about 3–4 km/h. The highest concentrations occur for wind speed between 5 and 10 km/h. This is consistent with emissions coming from an elevated source, high winds are needed to bring the emissions down to the ground. The Tung Chung site is much further from the airport than the LAX monitoring site and does not need winds as strong as those needed at LAX to bring emissions back to ground level. Also the Tung Chung monitoring site is located 21 m above the ground while the LAX monitor was near ground level, so the Tung Chung site requires much lower winds to bring the plume from the airport down to the level of the monitor.

The relationship of CO to wind direction and speed is seen in Fig. 8. A broad area of high CO is seen in the direction of the airport, with the highest concentrations of CO being in about the same region as the highest concentrations of SO_2. This area of high CO is likely a mixture of CO from ground vehicles going in and out of the airport and CO from aircraft. Traffic density around HKIA is much less that around LAX and ground vehicles do not dominate CO concentrations as at LAX. At high wind speeds, several patches of high CO are seen. These areas have high uncertainties and are not interpreted here. Some NO, NO_2, and NO_x data from Tung Chung showed internal inconsistencies that were not resolved therefore data for these gases were not part of this work.

Unlike the LAX data, Tung Chung data has hourly concentrations of RSP. The dependence of RSP on wind direction and speed is seen in Fig. 9. Once again, the impact of the airport is evident as some of the highest concentrations of RSP are found in the same region that also has high SO_2. Many of the highest values lie in regions at high wind speeds that have large uncertainties making it imprudent to attempt to interpret them. However, the peak at about 50° and 18 km/h has a signal-to-noise ratio of greater than 5 and, thus, may be related to a real source. Indeed, this direction points toward a major construction site for the Hong Kong West Rail that is about 18 km distant. The region from 0° to 60° and 5–10 km/h has relatively high RSP values. This may be the result of emissions from ships that pass through the channel to the north of Tung Chung.

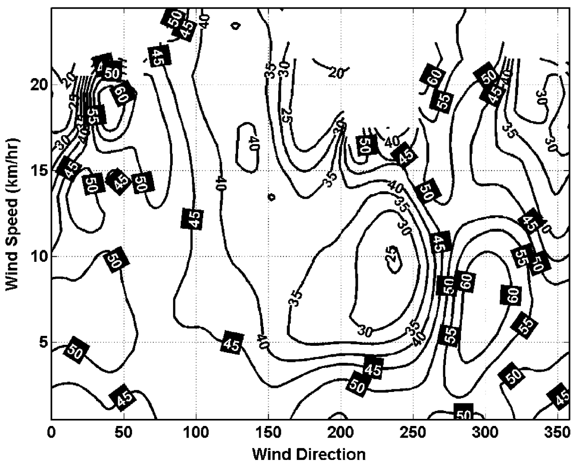

Fig. 9. Nonparametric regression Tung Chung RSP on wind speed and direction. The average RSP concentration contours are given in μg/m^3. Contours are not shown at points where the average RSP value is less than twice the estimated standard deviation. FWHM for wind direction is 25 and 4 for wind speed. Contours with high values are shown as white numbers on a black background.

5. Summary and conclusions

The impact of two major urban airports on local air quality has been studied using nonparametric regression of hourly pollutant concentrations on wind speed and direction. A major finding is that SO_2 can be used as a tracer for aircraft emissions at Los Angles International Airport and Hong Kong International Airport. While there are other sources of SO_2, these can be separated out based on wind direction and speed using nonparametric regression. It is likely that SO_2 can be used as tracer for aircraft emissions at other major airports. Thus, the concentrations of other pollutants in the same region of the nonparametric regression plots that have high concentrations of SO_2 can be attributed to aircraft. Based on this idea, CO and NO_x from ground vehicles dominated the emissions from aircraft at LAX. However, at Tung Chung near HKIA aircraft operations are a significant contributor to CO and RSP.

References

Colls, J., 2002. Air pollution: An Introduction, 2nd Edition, Spon Press, London, pp. 152–153.

Härdle, W., 1990. Applied Nonparametric Regression. Cambridge University Press, Cambridge.

Henry, R.C., Chang, Y.-S., Spiegelman, C.H., 2002. Locating nearby sources of air pollution by nonparametric regression of atmospheric concentrations on wind direction. Atmospheric Environment 36, 2237–2244.

Popp, P.J., Bishop, G.A., Sedman, D.H., 1999. Method for commercial aircraft nitric oxide emission measurements. Environmental Science and Technology 33, 1542–1544.

Spicer, C.W., Holdren, M.W., Smith, D.L., Hughes, D.P., Smith, M.D., 1992. Chemical composition of exhaust from aircraft turbine engines. Journal of Engineering for Gas Turbines and Power-Transactions of the ASME 114, 111–117.

Wand, M.P., Jones, M.C., 1995. Kernel Smoothing. Chapman and Hill, Ltd., New York and London.